Report of the Joint DOE/NSF

Ad Hoc Review Committee on

U.S. LHC Software and Computing

Glen Crawford, Chair

March 2000

EXECUTIVE SUMMARY

CMS and ATLAS will be large, general-purpose detectors used to observe very high-energy proton-proton collisions at the Large Hadron Collider (LHC). This facility is now under construction at CERN, the European Laboratory for Particle Physics near Geneva, Switzerland. In order to reap the scientific benefits of over $0.5 billion of U.S. investment in the LHC, the LHC software and computing projects must be successful in enabling physics analysis.

An *ad hoc* peer review of the U.S. LHC software and computing efforts was held on January 18-20, 2000, at the Germantown, MD offices of the Department of Energy (DOE). The primary purposes of this review were to assess the readiness of the collaborations to proceed to the next stage in their projects (including the formal setting of baselines later this year), to identify key areas which may need additional attention, and to help set priorities for near-term funding allocations. The expert reviewers provided comments during the review, both to the U.S. LHC collaborations and to the DOE and the National Science Foundation (NSF). These comments, and those provided afterward in letters from the reviewers, form the basis of this report.

The two U.S. collaborations have each proposed funding levels totaling $60 million to $76 million for their software and computing projects through FY-2005, followed by ongoing operations of about $20 million annually for each. Support at this level would have a substantial impact on the rest of U.S. high energy physics (HEP) program, which is already strapped for resources. Still, this funding level represents only a few percent of the total U.S. HEP program and will be a crucial component of the support of research for roughly one third of the U.S. HEP community. Whether the proposed funding levels are correct, and whether they can be achieved, remains to be determined.

The committee found that considerable progress has been made on software and computing by the U.S. LHC collaborations since the previous review of this activity in May 1999. The overall strategies of both the U.S. ATLAS and U.S. CMS software and computing projects were found to be sound. The U.S. groups have a strong competence in the areas of software and computing and they are making significant and important contributions.

Both projects will provide a mixture of deliverables to the U.S. user community (primarily access to data and computer power for analysis) and a well-defined set of software deliverables to the larger international community. In order to keep close coupling between the U.S. and international efforts, the committee stressed the importance of maintaining strong communication links not only within the ATLAS and CMS collaborations, but also between DOE, NSF, other funding agencies, and CERN.

The committee applauded both the U.S. ATLAS and U.S. CMS groups for taking strong leading roles in important areas of core software that will be used by the full collaborations. The immediate need for hiring software professionals to work on these key projects was stressed repeatedly by the members of the committee and was identified as the highest priority. Hiring professionals now will save significant resources in the long term, as they will produce much more coherent and maintainable structure and code than that produced by non-experts.

Present plans call for the U.S. CMS and ATLAS national ("Tier 1") computer facilities to be located at the Fermi and Brookhaven National Laboratories, respectively. Both projects estimate that the Tier 1 capacity will be approximately 20% of the corresponding facilities at CERN and assume a distributed model for computing and analysis of the data. Both Tier 1 plans were detailed and well-conceived and both experiments should begin staffing these facilities and procuring initial equipment resources to develop the Tier 1 sites for use by the U.S. LHC community. The committee members generally endorsed the distributed regional center ("Tier 2") concept and how it will interact with the Tier 1 center. However, scoping and planning for the Tier 2 centers are in a very preliminary phase. Uncertainties include the contributions from the host universities to these centers and the possibility of funding from other NSF programs.

The two U.S. efforts have proposed somewhat different management structures to carry out their software and computing efforts. These structures reflect the different styles and histories of the two collaborations and the personalities of their leadership, and both systems appear to be working well. From perspective of the funding agencies, the closer coupling of the software and computing efforts into U.S. ATLAS project management seems preferable to the more separate U.S. CMS arrangement. Drafts of the U.S. CMS and U.S. ATLAS Project Management Plans (PMPs) are well along, but will need some refinement over the coming months as the Agencies more precisely define the scope and funding profiles for these projects.

The near-term activities proposed by the two U.S. collaborations seem reasonable, and the U.S. LHC software and computing efforts will need a robust ramp-up in the near term in order to be ready for data taking in 2005. There appears to be an urgent need in FY2001 and perhaps beyond for more resources than the Agencies can currently provide for the U.S. LHC software and computing projects. The collaborations and their project managers should carefully consider how best to use their flexibility to optimize the total U.S. LHC efforts, particularly in FY2001.

The process of defining the appropriate scope of the U.S. LHC software and computing projects and the associated funding is on the critical path for establishing the baselines for these projects. This will necessarily be an iterative process and the Agencies must give it high priority if the baselines are to be established in a timely manner. Using a draft scope definition, the Agencies should consult with the U.S. LHC managers, and other managers of large computing projects, to discuss extrapolations to the U.S. LHC software and computing projects. DOE and host laboratory management also need to consult and discuss the appropriate level of resources that the laboratories should apply to the U.S. LHC projects out of their ongoing base programs.

Following these meetings, the Agencies should finalize the scope definitions and provide the collaborations with funding profiles, as was done for the U.S. LHC detector construction projects. The U.S. LHC software and computing project managers would then develop the detailed technical, cost and schedule baselines and hold internal reviews of them. In parallel, they would update and refine their Project Management Plans in preparation for a full baseline review about six months after the definition of the scope and funding profiles by the Agencies.

Table of Contents

Executive Summary

Appendices

1. INTRODUCTION

CMS and ATLAS will be large, general-purpose detectors used to observe very high-energy proton-proton collisions at the Large Hadron Collider (LHC) now under construction at CERN, the European Laboratory for Particle Physics near Geneva, Switzerland. The LHC will be the highest energy accelerator in the world for many years following its completion in 2005. It will provide two proton beams, circulating in opposite directions, at an energy of 7 trillion electron volts (TeV) each, almost an order of magnitude more energy than presently achieved at the Tevatron (1 TeV per beam), at Fermi National Accelerator Laboratory (Fermilab) outside Chicago.

The two large detectors will measure and record the results of the more interesting collisions. They will be among the largest and most complex devices for experimental research ever built, and the events that they see are expected to point to exciting, even revolutionary, advances in our understanding of matter and energy. The large increase in energy over that presently available may well lead to an understanding of the origin of mass and the discovery of new families of subatomic particles.

The U.S. scientific community strongly and repeatedly endorsed U.S. involvement in the LHC program. Numerous groups of U.S. scientists at universities and national laboratories, historically supported by both the Department of Energy (DOE) and the National Science Foundation (NSF), expressed great interest in the potential physics of the LHC and in 1994 they tentatively joined the international collaborations designing the CMS and ATLAS detectors. In 1996, DOE and NSF formed the Joint Oversight Group to coordinate and manage these efforts and to negotiate an appropriate U.S. role in the LHC program.

In December 1997, the heads of DOE, NSF and CERN signed an agreement on U.S. participation in the LHC program. This was further detailed by the Experiments and Accelerator Protocols signed later that month, committing the U.S. to spend a total of $531 million on LHC construction projects, with $200 million for aspects of the accelerator and the remainder supporting the efforts of the U.S. high energy physics (HEP) community in the construction of the two large detectors. The U.S. efforts on the detectors were formalized into construction projects with baselines established in 1998.

U.S. physicists are participating in many aspects of the detectors, including important management roles. With approximately 300 physicists from 30 U.S. universities and 3 national laboratories working on each of the two large detectors, the U.S. groups comprise roughly 20% of the full collaborations and the U.S. groups plan to provide a comparable portion of each detector.

As with past large detector projects, the LHC research program, including the computers and software needed for the physics data analysis, was not made part of the detector construction projects. However, the U.S. LHC research program must be successful if the U.S. HEP community is to reap the scientific benefits of the U.S. investment in the LHC. In addition, the international scientific community is depending on the U.S. to hold up its share of the collaborative effort. With the construction projects for both of the large general purpose detectors and the accelerator well underway, the Joint Oversight Group decided that it is now time for an assessment and formal organization of the U.S. LHC Research Program, including the software and computing projects that will be required to generate physics results over the life of the experiments.

The U.S. LHC Research Program will be a joint effort of DOE and NSF, utilizing the oversight structures established for the U.S. LHC Construction Project, as detailed in the DOE/NSF Memorandum of Understanding concerning U.S. participation in the LHC Program. In particular, in the coming year the first formal baseline reviews of the Software and Computing Projects of both U.S. ATLAS and U.S. CMS will be conducted, analogous to the DOE/NSF reviews of the U.S. ATLAS and U.S. CMS Detector Construction Projects.

To help the collaborations prepare for the formal baseline review, and to monitor progress on the existing U.S. R&D efforts directed towards LHC software and computing, the DOE and NSF held a joint *ad hoc* peer review of these efforts on January 18-20, 2000, at the Germantown, MD offices of DOE. The charge given to the reviewers is shown in Appendix A. The review committee was composed of experts in computing for high-energy physics and related fields, and the committee membership is detailed in Appendix B. A list of attendees at the review is given in Appendix C. Separate presentations on different days were made for the U.S. ATLAS and U.S. CMS computing projects, and a half-day was devoted to common projects.

After discussing the presentations, committee members provided feedback to the U.S. LHC collaborations on the same day, and then wrote formal review letters detailing their comments on various aspects of the software and computing projects. These letters and the discussions during the review provided the basis for this DOE/NSF report, written by the Chair of the review committee (a DOE staff member) in collaboration and consultation with committee members and Agency personnel.

This report and its recommendations represent the views of committee members on issues raised during the review, but it does not attempt to portray the personal opinions of every reviewer nor provide a comprehensive summary of all issues related to LHC computing efforts. It is intended as a compendium of expert advice to the funding agencies and the U.S. and international collaborators on the ATLAS and CMS experiments on how best to achieve the goals of the software and computing projects.

2. OVERALL SCOPE AND STRATEGY

The *ad hoc* review committee found that the overall strategies of both the U.S. ATLAS and U.S. CMS software and computing projects were sound. The U.S. groups have a strong competence in the areas of software and computing and they are making significant and important contributions. Both projects will provide a mixture of deliverables to the U.S. user community (primarily access to data and computer power for analysis) and a well-defined set of software deliverables to the larger international community of collaborators. This will allow the U.S. community to play a leading role in defining the software structure and also provides the platform for U.S. physicists to take an active role in all aspects of the physics output of the experiments.

Both collaborations stressed that providing U.S. users a "transparent" and effective means of remote collaboration was one of their key missions.

The committee members questioned the assumptions that went into defining the overall scope of the U.S. projects and hence the overall funding requirements. In particular, does the scope of the proposed U.S. effort correspond to the size of the U.S. participation in the collaborations and is it appropriate to meet the needs of the LHC experiments? The traditional "rule" of 1/3:2/3 contributions from CERN versus the participating collaborations was assumed to derive U.S. contributions, but several committee members questioned whether CERN would hold to this guideline. CERN representatives at the meeting indicated that this was their baseline assumption. Further, both U.S. collaborations assumed that their contributions to the computing requirements would go approximately equally to a large central computing resource (a "Tier 1" Facility, assumed to be about 20% of the "Tier 0" Facility at CERN) and the sum of distributed regional ("Tier 2") resources. Several committee members questioned the efficiency with which one would be able to marshal the latter resources. Finally, the long-range funding models assumed that resources needed after LHC turn-on in 2005 would be flat after correcting for inflation, whereas some committee members felt the overall resource needs (particularly for software support personnel) would decrease with time.

While the near-term activities proposed by the two U.S. collaborations seem reasonable, they will be limited by the funding that can be made available from the Agencies. The longer-term projections for manpower and equipment costs appear to be roughly appropriate, but detailed analyses should be made based on recent experience with similar projects.

The committee members found that it would be helpful for the Agencies to define which elements should be included within the two projects and their cost estimates. Each project presented their best analysis of the costs. However, the criteria for where to place the costs in the overall analysis were not consistent between projects. For example, the projects treated network costs in different manners and placed them in different locations in the overall cost estimates. DOE and NSF should provide clear direction to the experiments as to which items should be included in future project cost estimates.

Although pleased with the considerable recent progress and the near-term plans for ATLAS, the committee members felt strongly that their planned long-term computing milestones, including a Mock Data Challenge in 2003 and the Computing Technical Design Report in 2004, would be too late in the project cycle to have the appropriate positive impact on the computing effort. Committee members advocated accelerating the schedule for these milestones as much as possible so that ATLAS can make an early evaluation of its overall computing efforts and implement the necessary corrections.

RECOMMENDATIONS:

- **Provide clear Agency direction to the U.S. collaborations as to which items should be included in the scope and future cost estimates for the projects.**

- **Define the appropriate overall magnitude of U.S. LHC Software and Computing Projects using input from CERN about their expected contributions to the overall LHC computing effort, as well as guidance from the U.S. funding agencies.**

- **Advance the major ATLAS computing milestones for an initial Mock Data Challenge and the Computing Technical Design Report to 2002 and 2003, respectively.**

3. INTERNATIONAL ISSUES

The committee members observed that U.S. work in many key areas of software is driving the international efforts in both ATLAS and CMS. While the U.S. groups clearly should not get too far out in front of the international collaborations, the strength of the U.S. effort is a positive sign for the U.S. community and will better facilitate the goal of U.S. physicists playing an important role in physics analysis. In order to maintain close coupling between the U.S. and international efforts, the committee stressed the importance of maintaining strong communication links not only within the collaborations, but also between DOE, NSF, other funding agencies, and CERN.

As discussed in Section 2, the level of support from CERN for the computing efforts of the LHC experiments must be established. In addition, there is the question of common operating funds and a future agreement between CERN and the funding agencies as to the appropriate definitions and levels of funding needed once the experiments are operating.

The committee members found that the U.S. CMS computing effort appeared to be very well integrated with the larger CMS collaboration, and that International CMS had produced detailed computing plans that allow an understanding of the U.S. contributions in the context of the entire experiment. Committee members found the planned U.S. efforts to be appropriately matched to the U.S. participation in the experiment, within the present uncertainties of scope.

In the case of ATLAS, the computing efforts have been recently reorganized, so they lag behind CMS in defining their computing project and its deliverables. While the committee members recognized the difficulty of assembling a complete and detailed picture of ATLAS computing needs (including schedules, milestones, and work breakdown structures) on a short timescale, they exhorted the ATLAS collaboration to do so. Lacking such information, the committee found it hard to determine whether the levels of U.S. efforts were appropriate.

RECOMMENDATIONS:

♦ **U.S. ATLAS and U.S. CMS should continue to work closely with their international partners to help ensure the success of the LHC computing projects.**

♦ **Similarly, DOE and NSF should maintain close contact with their international partners (CERN and other funding agencies).**

♦ **The ATLAS collaboration should put together a complete and detailed picture of their computing needs and plans, including schedules, milestones, and work breakdown structure, as soon as possible. This is a prerequisite before U.S. ATLAS computing efforts can undergo a complete baseline review.**

4. USER FACILITIES

The U.S. CMS and ATLAS national (Tier 1) computer facilities will be located at Fermi and Brookhaven National Laboratories, respectively. At this time, there are large uncertainties in algorithm, software, and technology developments that will be needed by both facilities five years in the future. Both projects estimate that the Tier 1 capacity will be approximately 20% of the CERN Tier 0 Center and assume a distributed model for computing and analysis of the data. The plans of the two U.S. projects show similar staffing levels and relatively equivalent hardware.

The two U.S. collaborations presented similar preliminary design concepts for distributed facilities embodied in several Tier 2 Regional Centers. These conceptual designs include limited personnel and maintenance costs, and excellent connectivity to the Tier 1 Center and Data Grid. These centers were not as well defined as the Tier 1 Centers, but would serve as primary analysis sites for collaborators in the local geographic region and would have major simulation capabilities and onsite storage cache. The Tier 3 Centers at other laboratories and universities were briefly mentioned without design details.

Both U.S. CMS and U.S. ATLAS Tier 1 plans were well conceived and both experiments should begin staffing these facilities and procuring initial equipment resources to develop the Tier 1 sites for use by the U.S. LHC community. Committee members noted that the justification for the U.S. CMS Tier 1 plans was not as developed as for U.S. ATLAS, and would benefit from some additional work.

Because the start date for production running is several years away, the committee members felt it was too early to comment on specific technology choices. Nevertheless, both experiments presented cost estimates that the committee found to be realistic, based on the models that were presented and on today's equipment prices projected several years into the future. In general, both Tier 1 designs were deemed "conservative" in terms of reliance on extrapolation of proven technologies and costs.

Despite the long lead-time for the user facilities subprojects, many committee members expressed the strong opinion that dedicated R&D efforts and preliminary work on these facilities should begin soon. Staff with primary responsibility for these Tier 1 Centers need to be hired this calendar year to keep these projects on track. The appointment of permanent managers who will be in charge of the Brookhaven and Fermilab Tier 1 Centers was seen as a vital step.

The committee members generally endorsed the Tier 2 concept in terms of how it will interact with the Tier 1 Center. However, the actual detailed scoping and planning for the Tier 2 centers is in a very preliminary phase. Uncertainties include the contributions from the host universities to these centers and the possibility of funding from other NSF programs. It was also noted that at this preliminary stage, the number of planned Tier 2 Centers and the process for identifying their locations have not been finalized. The presentations assumed that the capability of each Tier 2 Center would be approximately 20% of a Tier 1 Center but there was no hard analysis presented to defend a Tier 2 model with five centers, rather than a model with 4 or 6 or 8 centers. As design and development efforts on the Tier 2 concepts progress, these uncertainties should be resolved.

There was also concern about the impact of various funding scenarios on the Tier 2 plans. Currently the majority of Tier 2 funding is expected to come from Information Technology initiatives in NSF, and some committee members wanted to hear about "fallback options" in case these funds were not forthcoming. Given the large current uncertainties in the Tier 2 designs, the collaborations had not yet directly addressed these issues, but asserted that the hierarchical computing model was flexible enough to accommodate such options, with some redirection of resources.

RECOMMENDATIONS

♦ **Find funding to allow staffing of the Tier 1 Centers to begin by the end of this calendar year at the latest, so that these facilities can begin to develop as central computing centers for each of the U.S. collaborations.**

♦ **Continue development of detailed plans for the Tier 2 Centers with a complete model of how these centers will actually operate and service the user community, including details of sizing, cost and functionality.**

5. USER INTERFACE AND PHYSICS ISSUES

The ATLAS and CMS collaborations have taken different approaches to software development and this has had a significant impact on the status of their software projects in general and user interface issues in particular. ATLAS chose to work on their detector and physics Technical Design Reports (TDRs) using a traditional FORTRAN environment, delaying the development of a future object-oriented (OO) framework. While this may have been the right choice from the point of view of detector optimization, it certainly puts more pressure on the collaboration now to design and implement a coherent OO software architecture quickly. At the same time they must maintain backwards-compatibility to the Monte Carlo simulations produced for the TDR milestone studies.

The U.S. ATLAS effort also specifically places a "Physics" subproject into their software organization, which is to maintain event generators and other pieces of physics infrastructure software. While the committee members generally approved of this structure and found physics analysis efforts strong and well-integrated into ATLAS, they stressed the need for near-future analysis tests (such as mock data challenges) to thoroughly test software releases and get user feedback. Further integration of users into the object-oriented code environment, through training and tutorials, was also encouraged.

The CMS collaboration chose to start development of an object-oriented framework for detector reconstruction and physics analysis early; their computing proposal was produced in 1996 and it gave a clear direction for software development. Detector and physics simulation software are now built into a framework of early OO software prototypes. On one hand, this situation makes ongoing detector optimization studies more painful, while on the other it provides crucial feedback for the development of core software. The committee members felt this development model would help ensure that CMS software will be ready when the first real data arrive. Nevertheless, CMS was clearly feeling the effects of user demands for high-performance computing now.

The committee heard of the difficulties in getting user resources in the U.S. for running detailed high-level trigger studies, and emphasized the importance of ramping up the Tier 1 user facility at Fermilab as soon as possible to facilitate those efforts. The ability of U.S. CMS members to run complete software packages at Fermilab and their local sites was viewed as a key element to the future success of the U.S. groups. The committee members encouraged a careful evaluation of the Fermilab contribution to the CMS computing efforts so that the U.S. CMS community is not "short-changed" by other programmatic concerns. The recent appointment of Jim Branson as U.S. CMS physics coordinator was viewed as a positive sign, and an important step towards the goal of enabling U.S. CMS groups to play a leading role in physics analysis.

RECOMMENDATIONS

- **Place a high priority on ramping up the capabilities of the Tier 1 user facilities and providing these resources to users.**

- **Continue efforts to train and fully integrate users into the software development process. Early adoption of mock data challenges and other tests of physics analysis goals are an essential component of this effort.**

- **Evaluate carefully the Fermilab contribution to the U.S. CMS computing efforts in the context of the overall U.S. program in high-energy physics.**

6. CORE SOFTWARE

The committee applauded both the U.S. ATLAS and U.S. CMS groups for taking strong leading roles in important areas of core software that will be used by the full collaborations. U.S. CMS has developed the fully object-oriented detector reconstruction package ORCA (Object Reconstruction for CMS Analysis), as well as a user interface and graphics suite, IGUANA (Interactive Graphical User ANAlysis). U.S. CMS is also leading database efforts in CMS, though the committee did not hear a specific report on this at the review. It was pointed out by some committee members that useful information with implications for the object database management system could be obtained from recent experience with the BaBar system.

Committee members singled out the CMS High-Level Trigger studies, which are making intensive use of the ORCA package, as a good example of the importance of testing software packages early in the development cycle, driven by end-user needs. While the committee members strongly endorsed this approach to software development, they also saw a need for U.S. CMS to develop a general quality assurance/quality control (QA/QC) plan for their software, complete with well-defined metrics that can be evaluated at each stage of development. They also proposed identifying a single person in CMS who will take charge and "ownership" of the QA/QC efforts. Similar efforts will also be needed for U.S. ATLAS.

U.S. ATLAS efforts are ramping up rapidly in concert with the reorganization of the overall ATLAS software effort, and the U.S. group proposes to take leading roles in defining the software framework and database structure. For the framework, they propose building on the software architecture model developed for the LHCb experiment in order to rapidly achieve a useable package for analysis, one that will evolve over time into the customized ATLAS software environment. The committee members strongly endorsed this collaborative effort. The committee members also expressed some concern about the coherence of the U.S. ATLAS database effort, especially given their limited resources. Some consolidation of this effort may be beneficial to the project.

The immediate need for hiring software professionals to work on these key projects was stressed repeatedly by the committee members and was identified as the highest-priority funding item for both U.S. ATLAS and U.S. CMS software and computing projects. All of the experts on the committee expressed the opinion that spending money to hire professionals now will save significant resources in the long term, as they will produce much more coherent and maintainable structure and code than that produced by non-experts.

RECOMMENDATIONS

- **Make additional funding for software professionals to contribute to both U.S. LHC core software efforts the highest software and computing priority for both collaborations.**

- **Develop general QA/QC plans for software, with well-defined metrics that can be evaluated at each stage of development, and identify a single person from each collaboration who will take charge of these efforts.**

- **Proceed to develop the U.S. ATLAS software framework in collaboration with LHCb, but consider consolidating U.S. ATLAS database efforts in light of resource limitations.**

7. COMMON PROJECTS

The committee heard presentations on a number of R&D efforts being worked on together by members of the two U.S. collaborations, as well common activities with others in the high energy physics community and elsewhere. These efforts include providing common infrastructure software, networking, collaborative tools, and large-area distributed computing services; many of these could become an integral part of the LHC computing efforts. In addition, the MONARC (Models of Networked Analysis at Regional Centers) project has provided important simulation tools to help develop the LHC computing model and evaluate different architecture scenarios. Groups in the U.S. have played an important role in MONARC and are continuing to use it for detailed optimization of their computing plans.

The committee members, while appreciative of the significant role MONARC has played in development of the computing models thus far, would have liked to see more validation of its predictions in scenarios where the scaling of computer resources needed to solve a particular problem is non-linear. A specific example is the case of BaBar, where such effects in database access have been experienced. Validation of the models on so-called "failure modes" would lead to a much better understanding of the critical components of a distributed computing system and hence a more robust computing solution. The committee members would also like to have seen more detailed MONARC studies related to Tier 2 centers, their optimal sizing, connectivity to Tier 1, etc. This would support the delineation of Tier 2 specifications as discussed in Section 4.

One common project that was not explicitly discussed as such, but identified by committee members, is the issue of networking and connectivity, both within the U.S. and from CERN to the U.S. The committee members observed that both projects hinge on the assumption of high-bandwidth connections (622 Mbits/sec) being available between the U.S. and CERN, and between the Tier 1 and Tier 2 centers, in 2005. The committee discussions indicated clearly that the required networks should be a common, “off-project” effort, not included in either U.S. ATLAS or U.S. CMS cost projections. However, both experiments must specify the bandwidth requirements they will need as a function of time and work with the funding agencies to assure that this level of connectivity is provided to the U.S. HEP community.

In general, the committee applauded the common efforts that had been undertaken so far and encouraged the U.S. collaborations to identify more areas in which they can collaborate. Many such possible projects were discussed during the review, including graphical user interfaces, configuration management tools, collaborative tools, and databases. While noting the difficulty of achieving success in these efforts in a large multinational collaboration, the committee felt the U.S. groups could set a strong example for their peers by embarking on new common projects.

RECOMMENDATIONS

- **Collaborate wherever possible to find common solutions to common problems in an efficient manner.**

- **Fund networking from a separate source as a common, "off-project" effort. Both U.S. ATLAS and U.S. CMS need to specify their bandwidth requirements in detail and work with the funding agencies to ensure their needs are met.**

- **Exploit the MONARC tools to undertake further simulation studies of the LHC computing model, specifically addressing "failure modes" where the scaling of resources becomes non-linear.**

- **Perform and verify detailed MONARC studies in support of the specifications of the Tier 2 design.**

8. FUNDING ISSUES

The U.S. LHC collaborations each proposed levels totaling $60 million to $76 million for their software and computing projects through FY2005, followed by ongoing operations of about $20 million annually for each. Support at this level of effort will have a substantial impact on the rest of U.S. high energy physics, which is already strapped for resources. Still, this effort represents only a few percent of the total U.S. HEP program and will be a necessary component of the support of research for perhaps one third of the U.S. HEP community.

The committee acknowledged that there were many uncertainties inherent in evaluating the long-term funding requests presented by U.S. ATLAS and U.S. CMS due to overall uncertainties in the project scale, scope, and the ever-changing nature of computing technology. To obtain a better evaluation of these issues, the committee members suggested that the Agencies request estimates of computing efforts currently underway for BaBar, CDF, D0, and the RHIC experiments, and how these might scale to the LHC experiments. The committee members also encouraged the U.S. LHC collaborations to provide the Agencies with a complete estimate of *all* the resources needed in their computing models, regardless of their presumed source of funding, so that a complete "roadmap" of the U.S. LHC computing efforts can be drawn up. Given these unresolved questions, the committee's conjecture was that the necessary level of funding was roughly consistent with what the collaborators proposed.

One aspect of the funding that was discussed in some detail was the question of the appropriate project contingency. The committee members generally agreed that contingency for a software and computing project was not *a priori* the same as for a construction project, and that the actual choice of a particular contingency percentage was highly uncertain. The question of what (if any) contingency to assign to these projects needs to be addressed via further guidance to the experiments from the Agencies. It was noted for the record that U.S. ATLAS and U.S. CMS used different contingency assumptions in their cost estimates.

The committee members were in general agreement that the near-term funding requests from both U.S. ATLAS and U.S. CMS were much easier to accurately evaluate, and many committee members strongly endorsed the proposed efforts of both groups as necessary for the eventual success of the software and computing projects. These efforts will require a robust ramp-up of funding in the near term in order to be ready for data taking in 2005. This ramp-up may well exceed the resources that the Agencies can provide for the U.S. LHC software and computing projects in FY2001 and perhaps beyond, given the long lead-times inherent to the Federal budget process. The committee members exhorted the Agencies to work hard to meet these needs, and advised the collaborations and their project managers to carefully consider how best to use their flexibility to optimize the total U.S. LHC efforts. For example, there are various "gray areas" (such as the support for detector-specific software and trigger studies) that have aspects of both detector construction and software/computing and could be reasonably funded from either project in the near term.

RECOMMENDATIONS

- **The Agencies should commission independent estimates of the long-term resources needed for the LHC software and computing projects based on the recent experiences of other HENP experiments, including BaBar, CDF, D0 and RHIC experiments.**

- **The U.S. LHC collaborations should provide the agencies with a complete estimate of *all* the resources needed in their computing models, regardless of their presumed source of funding**.

- **The near-term software and computing needs of both U.S. LHC groups must be met for the eventual success of the projects, and the U.S. groups and the Agencies should work together to meet them.**

9. PROJECT MANAGEMENT

Overall, both U.S. and International CMS are further along than ATLAS in software and computing effort and organization. U.S. CMS is well along in the selection and hiring process for its Level 1 Project Manager for Software and Computing and it is important to follow through and make this appointment as soon as possible. This person in turn should expeditiously appoint the Level 2 managers.

The committee praised the recent major structural decisions made for U.S. ATLAS computing. With most of its software and computing managers in place, the committee members felt U.S. ATLAS was in a good position to move forward quickly on the tasks at hand. There was some concern about the availability of some key people in the software management structure who have ongoing commitments to other projects. U.S. ATLAS management should work to ensure that these people can rapidly ramp-up to a full-time commitment to ATLAS.

The two U.S. efforts have proposed somewhat different management structures to carry out their software and computing efforts. There are two major differences:

1) The U.S. ATLAS Software and Computing Project has been brought into the existing management structure for the U.S. ATLAS detector construction project. The Associate Project Manager for Software and Computing reports to the U.S. ATLAS Project Manager and through him to the Brookhaven Laboratory management. The U.S. CMS Software and Computing Project Manager, on the other hand, reports directly to the Fermilab management in parallel to the U.S. CMS Project Manager for detector construction.

2) In U.S. ATLAS, the responsibilities of the Associate Project Manager include the software for physics and for reconstruction and detector subsystems, whereas for U.S. CMS these responsibilities remain directly with the U.S. CMS Collaboration, outside of both the detector and the software and computing projects.

These differences reflect the style and history of the two collaborations and the personalities of their leadership. In the case of the U.S. CMS collaboration, there also appears to be a concern to avoid being dominated by Fermilab. In principle, with the right people both systems can (and are) working well. Still, thought should be given to possible problems that could develop in the future and how best to organize the management structure to address any such problems. The closer coupling of the software and computing efforts into the U.S. ATLAS project management seems *a priori* preferable to the more separate U.S. CMS effort. In particular, having a single person directly responsible for U.S. CMS would facilitate management by the Agencies.

Both the U.S. CMS and U.S. ATLAS Project Management Plans (PMPs) are well along, but will need some refinements over the coming months as the Agencies more precisely define the scope and funding profiles for these projects. While the present drafts represent good starts, the two groups should be encouraged to borrow and combine the best features from the two drafts; maximum overlap and consistency should be encouraged where appropriate. For example, the thresholds for Change Control Boards should be made consistent, perhaps even with identical wording, and the lists of responsibilities of the various players might be strengthened by a judicious combination of the lists from the two drafts.

Some care needs to be taken in the definition of the responsibilities of the U.S. CMS Software and Computing Board (USSCB). While it is important that the Project Manager for Software and Computing stay in close communication with the U.S. CMS Collaboration and listen closely to its advice, he/she has direct responsibility for the project and the USSCB is not his/her boss. This point can be made clear by appropriate modifications to the draft PMP.

The Fermilab Directorate should in the near future establish an advisory structure external to the U.S. CMS project for the purpose of monitoring both the management and technical progress of the U.S. CMS Software and Computing Project. This advisory committee should report to a high level in the Fermilab Management (not to the Director of the Computing Division since the project is closely coupled to this division and being in charge of this advisory committee could put him in an awkward position with possible conflicts of interest).

DOE and host laboratory management need to consult and discuss the appropriate levels of resources that the laboratories should apply to the U.S. LHC projects out of the ongoing laboratory base programs. These projects will become an increasingly important component of the scientific programs at these laboratories. Even though the data may physically be collected elsewhere, the scientific strength and reputation of the host laboratories will greatly depend on the success of the U.S. LHC efforts and the critical roles played by the laboratories in that success. It is therefore reasonable that some effort be redirected by the laboratories to these important projects.

The process of defining the scope of the U.S. LHC software and computing projects and the associated funding is on the critical path leading to establishment of the baselines for these projects. This will necessarily be an iterative process and the Agencies must give it high priority if the baselines are to be established in a timely manner. The process will involve the Agencies working both internally and with the collaborations. Consultation with the broader HEP community may also be useful.

A suggested way to proceed would be for the Agencies to establish a draft project scope, with explicit assumptions on resources to be expected from elsewhere. Using this draft scope definition, project managers from the other recent and ongoing HENP computing projects would then be requested to estimate the corresponding time and resources used for their projects and to critique the adequacy of those resources. The Agencies would then convene a meeting with these managers together with the U.S. LHC managers to discuss the implications for the U.S. LHC software and computing projects.

Following this meeting, the Agencies would finalize the scope definitions and provide the collaborations with funding profiles (including guidance as to the appropriate levels of contingency), as was done for the U.S. LHC detector construction projects. The U.S. LHC software and computing project managers would then have the detailed technical, cost and schedule baselines developed and hold internal reviews ("take ownership") of them. In parallel, they would update and refine their Project Management Plans in preparation for a technical, cost, schedule and management baseline review about 6 months after the definition of the scope and funding profiles by the Agencies. This review should result in the formal establishment of these baselines.

RECOMMENDATIONS

- **U.S. CMS should expeditiously appoint the permanent Level 1 and 2 project managers for its software and computing project. U.S. ATLAS should similarly appoint the deputy manager of its Tier 1 user facility.**

- **U.S. CMS should consider closer integration of the management of the detector construction project and the software and computing project, with a single person responsible overall.**

- **U.S. CMS and U.S. ATLAS should refine their draft Project Management Plans, borrowing from one another and making the two plans consistent with one another to the extent possible.**

- **The Fermilab Directorate should in the near future establish an advisory structure external to the U.S. CMS project for the purpose of monitoring both the management and technical progress of the U.S. CMS Software and Computing Project.**

- **DOE and the host laboratory management should discuss and decide the appropriate levels of resources that the laboratories should apply to the U.S. LHC software and detector projects out of the ongoing laboratory base programs.**

- **The Agencies should in the near future define the scope of the U.S. LHC software and computing projects, including explicit assumptions on the resources to be expected from elsewhere and guidance on determining the appropriate project contingency.**

- **The Agencies should use this project scope definition, along with data from similar recent software and computing projects, to determine the funding profiles to be expected for the U.S. LHC software and computing projects.**

- **U.S. LHC management should use this scope definition and funding profile to propose baselines and iterate Project Management Plans in preparation for an Office of Science baseline review six months later.**

10. ACTION ITEMS

- **DOE/NSF to supply software and computing project scope definition and funding profile as soon as practicable.**

- **U.S. LHC management to use this scope definition and funding profile to propose baselines and iterate Project Management Plans in preparation for an Office of Science baseline review six months later.**

APPENDICES

A. Charge to the Review Committee

The LHC at CERN is scheduled for completion in 2005, and the Construction Projects for both detectors (ATLAS and CMS) and the accelerator are well underway. Now is an appropriate time to begin seriously assessing aspects of the LHC Research Program, including the Software and Computing Projects which will be required to generate physics results for the life of the experiments. The U.S. LHC Research Program will be a joint effort of the DOE and NSF utilizing the oversight structures established for the U.S. LHC Construction Project, as detailed in the DOE/NSF Memorandum of Understanding concerning U.S. participation in the LHC Program. In the past, the funding agencies have responded to the U.S. collaborators' requests for R&D funds for software and computing in a mostly *ad hoc* manner.

As the effort moves into the pre-operations phase it is necessary to formalize the arrangements for funding and oversight**.** In particular, in the coming year we plan to conduct the first formal Technical, Cost, Schedule and Management Baseline Review of the Software and Computing Projects of both U.S. ATLAS and U.S. CMS, analogous to the annual DOE/NSF reviews of the U.S. ATLAS and U.S. CMS Detector Construction Projects.

To help the collaborations prepare for the formal Baseline Review, and to monitor progress on the existing U.S. R&D efforts directed towards LHC Software and Computing, we are conducting a peer review of these efforts on January 18-20, 2000, in Washington, D.C. We are also asking the collaborations to identify common areas between the two U.S. Software and Computing Projects which may be reasonably combined to achieve cost or schedule efficiency, and to highlight these common areas during the January review.

The primary purpose of this review is to assess the collaborations' readiness to proceed to the next stage in their projects and to identify key areas which may need additional attention. Specifically, the review committee should evaluate:

1) The overall strategy and scope of the U.S. software and computing efforts, and their relationship to the plans of the international community;
2) The proposed designs of the U.S. ATLAS and U.S. CMS computing facilities;
3) The realism of the proposed schedules*;*
4) The adequacy of the long-term funding profiles proposed by the collaborations;
5) The commonalities between the U.S. ATLAS and U.S. CMS software and computing plans and the experiments' plans to seek common approaches to common problems*;*
6) The appropriateness of the management structures and the Project Management Plans presented by the collaborations; and
7) The schedules of work and cost estimates for the coming year.

In addition, the committee is asked to comment on relative priorities of the various elements of the software and computing projects.

B. Committee Membership

Name	Affiliation(s)
Bruce Allen	U. Wisconsin
Ian Bird	TJNAF
Glen Crawford, Chair	DOE
Bob Diebold	consultant
Patrick Dreher	MIT
Bruce Gibbard	BNL (ATLAS)
Nick Hadley	U. Maryland
Matthias Kasemann	FNAL (CMS)
Richard Mount	SLAC
Tomasz Skwarnicki	Syracuse Univ.
Terry Schalk	UC Santa Cruz
Walter Toki	Colorado State U.

C. List of Attendees

US ATLAS
John Huth
Bruce Gibbard
Howard Gordon
Tom Kirk
Craig Tull
Jim Shank
Frank Merritt
David Malon
Larry Price
Torre Wenaus
Ian Hinchliffe
Krzys Sliwa
David Quarrie

US CMS
Matthias Kasemann
Joel Butler
Ken Stanfield
Vivian O'Dell
Harvey Newman
Lucas Taylor
David Stickland
Darin Acosta
Dan Green
Paul Avery
Winston Ko
Jim Branson

Computing Grids
Ian Foster, ANL

CERN/International
Hans Hoffmann, CERN
Norman McCubbin, ATLAS
Martti Pimia, CMS

Reviewers
Glen Crawford, Chair
Bruce Allen
Ian Bird
Bob Diebold
Patrick Dreher
Nick Hadley
Richard Mount
Tomasz Skwarnicki
Terry Schalk
Walter Toki

DOE observers
PK Williams
Tim Toohig
Jack Ritchie
Dan Lehman
Lowell Ely
Dan Hitchcock
Jim Yeck
John O'Fallon
Peter Rosen
Vicky White

NSF observers
Jack Lightbody
Alex Firestone